

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problems Mailbox.**

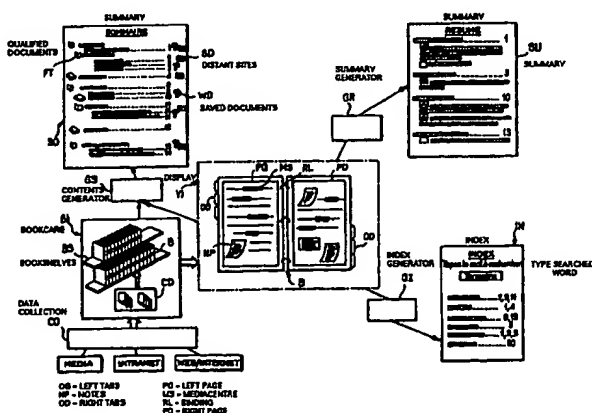


DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

(51) Classification internationale des brevets ⁶ : G06F 3/033, 17/30	A1	(11) Numéro de publication internationale: WO 98/55914 (43) Date de publication internationale: 10 décembre 1998 (10.12.98)
<p>(21) Numéro de la demande internationale: PCT/FR97/00983</p> <p>(22) Date de dépôt international: 3 juin 1997 (03.06.97)</p> <p>(71) Déposant (pour tous les Etats désignés sauf US): TEREIS CORPORATION [US/US]; 1001 Marshall Street, Redwood City, CA 94063 (US).</p> <p>(72) Inventeurs; et (75) Inventeurs/Déposants (US seulement): SAFARS, Michel [FR/FR]; 2, allée Elisabeth de Feydeau, F-91190 Gif sur Yvette (FR). FRACHON, Fabrice [FR/FR]; 180, avenue du Général Leclerc, F-91190 Gif sur Yvette (FR). SYLVESTRE DE SACY NOUGARET DURAFORT, Pierre [FR/FR]; 4, rue Alfred Stevens, F-75009 Paris (FR). VERRECCHIA, Thierry [FR/FR]; 23, rue de Neuville, F-95000 Cergy (FR).</p> <p>(74) Mandataires: ALLANO, Sylvain etc.; Pontet Allano & Associés S.E.L.A.R.L., 25, rue Jean Rostand, Parc-Club Orsay-Université, F-91893 Orsay Cedex (FR).</p>	<p>(81) Etats désignés: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, brevet ARIPO (GH, KE, LS, MW, SD, SZ, UG), brevet eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</p> <p>Publiée Avec rapport de recherche internationale.</p>	

(54) Title: METHOD AND SYSTEM FOR DATA QUALIFICATION IN PRE-COLLECTED DOCUMENTS

(54) Titre: PROCÉDE ET SYSTEME DE QUALIFICATION DE DONNEES AU SEIN DE DOCUMENTS PREALABLEMENT COLLECTES



(57) Abstract

The invention concerns a method for data qualification in pre-collected documents, in particular from sites connected to a communication network, said document being consulted in the form of a book pages whereof they constitute the contents. Data or information can be selected within each page of the document being consulted and contained in a book, so as to generate from said selected data tools for accessing the book contents, in particular an index, summary or synopsis. The invention is useful in particular for off-line consultation of documents collected on the Web.

(57) Abrégé

Procédé pour qualifier des données au sein de documents préalablement collectés, notamment à partir de sites connectés à un réseau de communication, ces documents pouvant être consultés sous la forme de pages d'un livre dont ils constituent le contenu. Des données ou informations peuvent être sélectionnées au sein de chaque page d'un document en cours de consultation et contenu dans un livre, de façon à générer à partir de ces données sélectionnées des outils d'accès au contenu de ce livre, notamment un index, un sommaire ou un résumé. Utilisation notamment pour des consultations hors-ligne de documents collectés sur le Web.

UNIQUEMENT A TITRE D'INFORMATION

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AL	Albanie	ES	Espagne	LS	Lesotho	SI	Slovénie
AM	Arménie	FI	Finlande	LT	Lituanie	SK	Slovaquie
AT	Autriche	FR	France	LU	Luxembourg	SN	Sénégal
AU	Australie	GA	Gabon	LV	Lettonie	SZ	Swaziland
AZ	Azerbaïdjan	GB	Royaume-Uni	MC	Monaco	TD	Tchad
BA	Bosnie-Herzégovine	GE	Géorgie	MD	République de Moldova	TG	Togo
BB	Barbade	GH	Ghana	MG	Madagascar	TJ	Tadjikistan
BE	Belgique	GN	Guinée	MK	Ex-République yougoslave de Macédoine	TM	Turkménistan
BF	Burkina Faso	GR	Grèce	ML	Mali	TR	Turquie
BG	Bulgarie	HU	Hongrie	MN	Mongolie	TT	Trinité-et-Tobago
BJ	Bénin	IE	Irlande	MR	Mauritanie	UA	Ukraine
BR	B Brésil	IL	Israël	MW	Malawi	UG	Ouganda
BY	Bélarus	IS	Islande	MX	Mexique	US	Etats-Unis d'Amérique
CA	Canada	IT	Italie	NE	Niger	UZ	Ouzbékistan
CF	République centrafricaine	JP	Japon	NL	Pays-Bas	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norvège	YU	Yougoslavie
CH	Suisse	KG	Kirghizistan	NZ	Nouvelle-Zélande	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	République populaire démocratique de Corée	PL	Pologne		
CM	Cameroun	KR	République de Corée	PT	Portugal		
CN	Chine	KZ	Kazakhstan	RO	Roumanie		
CU	Cuba	LC	Sainte-Lucie	RU	Fédération de Russie		
CZ	République tchèque	LI	Liechtenstein	SD	Soudan		
DE	Allemagne	LK	Sri Lanka	SE	Suède		
DK	Danemark	LR	Libéria	SG	Singapour		
EE	Estonie						

- 1 -

**"PROCÉDE ET SYSTEME DE QUALIFICATION DE DONNEES AU SEIN
DE DOCUMENTS PREALABLEMENT COLLECTES"**

La présente invention concerne un procédé de qualification de données au sein de documents
5 préalablement collectés. Elle vise également un système pour sa mise en oeuvre.

Le développement des supports d'informations et des réseaux de communications tels qu'Internet a conduit à une augmentation considérable des masses de données et
10 d'informations qui peuvent être consultées et collectées. Les utilisateurs de ces nouvelles techniques se trouvent rapidement submergés par des volumes de données qu'il devient indispensable de ranger, de classer et de qualifier, tout particulièrement lorsqu'on souhaite
15 effectuer des consultations hors-ligne de ces données. Actuellement, si les internautes chevronnés peuvent utiliser tant bien que mal des techniques personnelles de classement des milliers de pages HTML collectées, il en va tout autrement pour les nouveaux utilisateurs
20 d'Internet qui accèdent à un univers de données qui leur est inconnu et ne possèdent en aucun cas la dextérité et les "tours de main" des internautes précités.

Le but de la présente invention est précisément de proposer un procédé de qualification de données qui
25 permette une réalisation simple et rapide de livres à partir d'ensembles de documents collectés à partir de diverses sources d'informations.

Ceci est obtenu avec un procédé pour qualifier des données au sein de documents préalablement collectés,
30 notamment à partir de sites connectés à un réseau de communication, ces documents pouvant être consultés sous la forme de pages d'un livre dont ils constituent le contenu. Il est à noter qu'il existe déjà un produit logiciel dénommé **Lois** (navigateur Web) qui intègre une

- 2 -

présentation de documents sous la forme de pages d'un livre.

Suivant l'invention, des données peuvent être sélectionnées au sein de chaque page d'un document en cours de consultation et contenu dans un livre, de façon à générer à partir desdites données sélectionnées des outils d'accès au contenu de ce livre.

On dispose ainsi avec le procédé de qualification selon l'invention d'un mode d'organisation et de gestion des données qui va bien au delà d'une simple compilation de pages consultées puisqu'il va devenir possible de traiter l'ensemble des données contenues dans les documents constituant un livre sous une forme globale grâce à des outils d'accès configurés pour chaque livre.

Par exemple, au titre des outils d'accès, le procédé selon l'invention peut comprendre une génération d'un index à partir des données mises en évidence dans un document, cet index signalant l'ensemble des pages dans lesquelles chacune des informations apparaît.

Cet index peut notamment être généré à partir d'informations surlignées, mais également à partir d'informations signalées par un traitement spécifique des caractères les composant. Il peut comporter une zone pour saisir un terme à rechercher dans l'index.

Un outil d'accès au contenu des documents consiste en un résumé généré à partir d'informations sélectionnées au sein d'un document. Ce résumé peut aussi être alimenté à partir de notes prises dans d'autres documents d'un livre.

Le procédé de qualification selon l'invention peut aussi avantageusement prévoir d'associer à un document une ou plusieurs notes apposées sur ce document à la demande d'un utilisateur. Ces notes se présentent par exemple sous la forme d'une feuille autocollante à apposer sur un document, dans laquelle l'utilisateur peut

- 3 -

saisir une portion de texte. Ces notes peuvent avoir la propriété d'être transparentes, permettant ainsi de visualiser le contenu de la note sans pour autant masquer le document. La recherche dans l'index peut d'ailleurs
5 aussi porter sur ces notes apposées.

Dans une version avantageuse du procédé selon l'invention, il est fourni un sommaire généré à partir de l'ensemble des documents consultés. Chaque document consulté est ainsi ajouté au sommaire de façon
10 chronologique, une réorganisation du sommaire de façon hiérarchique pouvant aussi être envisagée. En particulier, lorsqu'il s'agit de documents collectés sur un réseau de communication tel qu'Internet, la réorganisation du sommaire peut être réalisée suivant
15 l'arborescence des adresses URL.

Le procédé selon l'invention peut en outre avantageusement comprendre une possibilité pour un utilisateur de déplacer l'ensemble des liens du sommaire de façon à générer sa propre arborescence, ainsi qu'un
20 masquage d'une ou plusieurs branches de cette arborescence.

Un signalement par le sommaire de documents modifiés sur des sites distants, de documents sauvegardés localement ou de documents qualifiés (annotations,
25 passages surlignés dans le document), peut en outre être prévu. En pratique, le procédé selon l'invention comprend en outre une génération d'un dossier contenant toutes les pages consultées.

Suivant un autre aspect de l'invention, il est
30 proposé un système pour qualifier des données au sein de documents préalablement collectés, notamment en visitant des sites connectés à un réseau de communication, mettant en œuvre le procédé selon l'invention, associé à des moyens pour collecter des documents, des moyens pour
35 stocker ces documents collectés, des moyens pour traiter

- 4 -

ces documents collectés et stockés et à des moyens pour visualiser et consulter le contenu de ces documents sous la forme de pages d'un livre, caractérisé en ce qu'il comprend des moyens pour sélectionner des données au sein
5 des pages consultées. Ce système comprend en outre des moyens pour générer des outils d'accès au contenu des documents stockés au sein d'un livre, notamment un index, un sommaire et un résumé.

Une page blanche personnelle peut être insérée à la
10 position courante, afin d'accueillir des notes textuelles ou graphiques ou tout autre élément. On peut d'ailleurs prévoir la possibilité de mettre en œuvre un logiciel de PAO pour cette page personnelle.

D'autres particularités et avantages de l'invention
15 apparaîtront encore dans la description ci-après. Aux dessins annexés donnés à titre d'exemples non limitatifs:

- la figure 1 représente un premier exemple de visualisation de documents en format "livre" par le procédé selon l'invention;
- 20 - la figure 2 représente un second exemple de visualisation de documents en format "calepin" par le procédé selon l'invention;
- la figure 3 représente un troisième exemple de visualisation de documents en format "bloc notes" par le
25 procédé selon l'invention;
- la figure 4A représente la structure d'une base de données mise en œuvre avec le procédé selon l'invention;
- la figure 4B représente les modes d'accès aux données dans une base de données mise en œuvre avec le
30 procédé selon l'invention;
- la figure 5 est une vue synoptique des fonctions essentielles procurées par le procédé selon l'invention.

On va maintenant décrire des exemples de visualisation de documents obtenus avec le procédé de

- 5 -

qualification selon l'invention, en référence aux figures 1 à 3.

L'écran 1 comprend deux pages respectivement gauche et droite PG, PD articulée autour d'une reliure RL, et un ensemble d'icônes disposés par exemple sur la partie supérieure de l'écran. Le livre B1 ainsi visualisé comporte un premier ensemble O11 d'onglets disposés sur son côté gauche et un second ensemble O12 d'onglets disposés sur son côté droit. Il va de soi qu'au gré des pages tournées, les nombres respectifs d'onglets situés instantanément respectivement à gauche et à droite sont variables. Le premier ensemble O11 d'onglets comprend un onglet OS "Sommaire" situé de préférence au sommet de ce premier ensemble d'onglets et un onglet marque-page MP, tandis que le second ensemble O12 d'onglets comprend un onglet OI "Index" situé de préférence au bas de ce second ensemble d'onglets. Les pages de gauche et de droite PG, PD reçoivent respectivement des documents actifs DG, DD. L'écran 1 comprend en outre une barre de navigation BN située par exemple sur sa partie supérieure et un ensemble d'icônes de commande EI. Il est en outre pourvu de zones de raccourci RP, RS pour atteindre respectivement les pages suivantes et précédentes.

Dans le second exemple de visualisation illustré par la figure 2, l'écran 2 comprend une première page supérieure PS et une seconde page inférieure PI contenant chacune des documents actifs DS, DI en cours de visualisation. Le calepin B2 ainsi visualisé comprend un premier ensemble d'onglets O21 situés sur le bord supérieur de la partie supérieure PS et un second ensemble d'onglets O22 situés sur le bord inférieur de la partie inférieure PI. L'ensemble d'icônes EI est par exemple disposé sur le côté droit de l'écran 2. Le premier ensemble d'onglets O21 comprend par exemple un

- 6 -

onglet "sommaire" OS, tandis que le second ensemble d'onglets O22 comprend par exemple un onglet "index" OI.

On peut également utiliser un mode de visualisation de type bloc-notes en référence à la figure 3, dans lequel un écran 3 comprend un ensemble d'onglets O31 disposé par exemple sur le bord supérieur d'une page unique PA contenant un document actif DA en cours de visualisation. L'ensemble d'onglets O31 comprend par exemple un onglet "sommaire" OS, des onglets intercalaires O2-O5 et un onglet d'index OI.

Une base de données DB mise en œuvre avec le procédé selon l'invention comprend, en référence aux figures 4A et 4B, un centre de documentation MS, une bibliothèque ("bookcase") BC, des étagères BS, des livres BO, des termes indexés TI, des liens LI, des données DA, des marques de livre BM, des notes NO et des surligneurs HI. Elle comprend en outre des drapeaux FL, un fichier d'accès AC, des drapeaux LS pour déterminer des niveaux de sécurité, une table de temps TT, un journal d'événements LO, des utilisateurs US, des groupes d'utilisateurs GR, des relations entre groupes et utilisateurs GU et un identifiant d'accès unique OA.

Le Centre de média ou de documentation MS, appelé aussi dans la suite MediaCenter, est le paradigme logique d'un centre de documentation. Il permet d'identifier les grandes sources d'informations. A titre d'exemple non limitatif, au sein d'une entreprise, il est possible d'implanter plusieurs serveurs de bases de données qui peuvent être ainsi consultés indifféremment. Le MediaCenter regroupe l'ensemble des bibliothèques appelé dans la suite BookCase.

Le BookCase est un espace de rangement pour un grand nombre de livres regroupés logiquement par un utilisateur. C'est le paradigme du meuble que représente une bibliothèque. Il est associé à un unique MediaCenter.

- 7 -

Le BookShelf est le paradigme d'une étagère. Il permet de regrouper thématiquement un ensemble de livres. Il n'existe aucune limite logique du nombre de livres. Seule la représentation à l'écran d'une étagère dans la
5 bibliothèque peut limiter leur nombre.

Les livres sont rangés dans des étagères et peuvent être déplacés dans n'importe lequel des éléments hiérarchiques précités.

Les MediaCenters peuvent être des éléments
10 partageables sur le Web. Un utilisateur navigant sur Internet peut décider de consulter une bibliothèque d'information. Il peut ainsi passer d'un MediaCenter dans un autre MediaCenter.

L'ensemble des mots indexés peut être regroupé sous
15 la forme d'un livre général des index afin de permettre une recherche sur l'ensembles des livres.

On va maintenant décrire des exemples de gestion des droits d'accès aux informations dans le cadre du procédé selon l'invention. On définit un ensemble de services
20 permettant l'administration des droits d'accès aux informations.

Chaque document possède une liste de lecteurs et de groupes de lecteurs ayant des droits spécifiques sur ce document. Les ressources possèdent des droits attribués
25 aux utilisateurs. Des droits peuvent être associés à:

- des centres de documentation (MediaCenter),
- des bibliothèques (Bookcase),
- des groupes de livres ou étagères (Bookshelf),
- des livres,
- 30 - des sections ou chapitres,
- des documents (ou liens) d'un livre.

Les droits suivants peuvent être appliqués:

- lecture,
- écriture,

- 8 -

- ajout (d'un livre pour des groupes de livres, d'un document pour des livres),

- appropriation (d'un groupe de livres, d'un livre, d'un document),

5 - backup,

- plages horaires (Timetable).

Un utilisateur autorisé à supprimer un groupe de livres ne peut pas pour autant supprimer les documents d'un livre, sauf s'il lui a été explicitement donné ce
10 droit.

La propriété d'un élément n'implique pas l'ensemble des droits disponibles mais la possibilité de modifier n'importe quel droit de cet élément. Un administrateur a un droit d'appropriation sur tous les éléments de la
15 bibliothèque. La notion de propriété permet d'identifier n'importe quel utilisateur s'étant approprié un document.

Le droit de backup n'implique pas le droit de lecture: un opérateur spécifique peut détenir ce droit sans pour autant être autorisé à consulter les données.

20 Les horaires permettent de limiter l'accès à des informations, par exemple par analogie avec les horaires d'ouverture d'une bibliothèque.

Seuls des opérateurs créés par l'administrateur ont la possibilité de créer des lecteurs et des groupes de
25 lecteurs. On peut définir pour les opérateurs les droits suivants:

- ajout d'utilisateurs

- suppression d'utilisateurs,

- suppression de lecteurs qu'ils ont créés,

30 - suppression de groupes de lecteurs qu'ils ont créés,

- suppression de lecteurs,

- suppression de groupes de lecteurs.

- 9 -

Il existe un groupe administrateur dont les membres ont les mêmes droits que l'administrateur. L'administrateur ne peut jamais être supprimé.

On peut également prévoir d'attribuer une durée de vie aux données qui est répercutée au travers du champ ExpirationDate de la base de données. L'introduction de cette durée de vie permet de résoudre des problèmes de confidentialité, de capacité de stockage et de conservation des données pertinentes. Elle renforce également le paradigme du livre étant donné que les lecteurs sont habitués à lire des quotidiens, des hebdomadaires, des mensuels.

Les utilisateurs et les groupes d'utilisateurs possèdent également une durée de vie permettant de restreindre pendant à certaine durée les droits de certains utilisateurs. Par ailleurs, il est possible de fixer une date d'expiration des mots de passe des utilisateurs pour les forcer à renouveler leurs mots de passe et accroître ainsi la sécurité.

On va maintenant décrire les fonctions essentielles procurées par le procédé selon l'invention, en référence à la figure 5. Des informations sont collectées (CO) sur le Web/Internet, sur un réseau Intranet ou sur des supports d'informations tels que des CD-ROM. Ces informations ou documents sont ensuite classés (CD) pour être rangés dans des livres B disposés sur des étagères BS, au sein d'une bibliothèque BI. Lorsqu'un utilisateur souhaite consulter un document, le livre B contenant ce document est extrait de la bibliothèque et est ouvert afin de permettre une visualisation VI des pages de ce livre. Ainsi, on peut visualiser une page gauche PG et une page droite PD pouvant tourner autour d'une reliure RL, ces pages étant éventuellement pourvues d'onglets OG, OD. Sur les pages des documents, l'utilisateur peut à son

- 10 -

gré surligner (MS) des mots ou des extraits du document, ou apposer des notes NP.

Au cours de la constitution du livre, le procédé selon l'invention inclut les fonctions suivantes:

- 5 - génération GS d'un sommaire SO;
- génération GR d'un résumé SU;
- génération GI d'un index IN.

Le surligneur a pour fonction de permettre une mise en relief d'informations par l'utilisateur. Ainsi, en
10 consultant à nouveau un document, il peut voir plus facilement ce qui l'intéresse. Cette fonction permet également de faire partager des informations clés à d'autres utilisateurs, notamment dans une architecture Intranet. Les mots surlignés servent aussi à la création
15 d'un index.

Le résumé est un document généré à partir d'éléments surlignés et de notes prises dans d'autres documents du livre.

Le sommaire est généré à partir de l'ensemble des
20 documents consultés. Chaque document consulté est ajouté au sommaire de façon chronologique. Un utilisateur peut accéder au sommaire à tout moment de la consultation. Il peut demander la réorganisation du sommaire de façon hiérarchique, par exemple suivant l'arborescence des
25 adresses URL dans le cas de documents prélevés sur Internet. L'ensemble des liens du sommaire peut être déplacé par l'utilisateur de manière à générer sa propre arborescence. Toute branche de cette arborescence peut être masquée. Le sommaire est capable de signaler à
30 l'utilisateur les documents modifiés sur des sites distants. Un dossier spécial appelé "pages non classées" contient toutes les pages consultées.

Chaque document téléchargé et appartenant au sommaire est analysé pour en extraire des mots clés. Ces mots clés
35 permettent de générer un index. Ils sont mis en relief

- 11 -

dans un document soit par des fonctions de caractère (Gras, Italique,...) ou en utilisant la fonction de surlignage.

L'index signale l'ensemble des pages où chacun des
5 mots clés apparaît, même s'il a été mis en relief dans un seul document.

L'index dispose d'une zone de saisie qui permet de se placer instantanément sur le terme recherché dans l'index. Cette zone de saisie peut également permettre
10 une recherche plus complète dans l'ensemble du livre. Dans ce cas là, un nouveau document contenant le résultat de la recherche est ajouté au livre.

Un utilisateur peut associer à un document une ou plusieurs notes qui sont positionnées au gré de
15 l'utilisateur sur le document. Lors d'une recherche dans l'index, il est possible de spécifier que la recherche s'effectue aussi dans les notes, ou bien uniquement dans les notes. Les notes se présentent sous la forme d'une feuille autocollante dans laquelle l'utilisateur peut
20 saisir une portion de texte.

Il est à noter que le procédé de qualification selon l'invention peut être associé à tout procédé d'édition et d'impression permettant de réaliser physiquement de vrais livres, calepins ou bloc-notes ou tout autre support
25 matériel. Par ailleurs, le procédé de qualification selon l'invention peut prendre en compte toute applette applicative disponible dans le domaine du multimédia.

Bien sûr, l'invention n'est pas limitée aux exemples qui viennent d'être décrits et de nombreux aménagements
30 peuvent être apportés à ces exemples sans sortir du cadre de l'invention. Ainsi, d'autres outils que ceux qui viennent d'être décrits peuvent être mis en œuvre pour la qualification des données. Il est à noter que les données et documents concernés par la présente invention peuvent
35 être d'origine quelconque et avoir été collectées sur des

- 12 -

réseaux de communications ouverts tels qu'Internet ou
privés (Intranet), ou encore sur des supports
d'informations de toute nature (CD-ROM, DVD-ROM, etc..).

- 13 -

REVENDICATIONS

1. Procédé pour qualifier des données au sein de documents préalablement collectés, notamment à partir de sites connectés à un réseau de communication, ces documents pouvant être consultés sous la forme de pages d'un livre dont ils constituent le contenu, caractérisé en ce que des données peuvent être sélectionnées au sein de chaque page d'un document en cours de consultation et contenu dans un livre, de façon à générer à partir desdites données sélectionnés des outils d'accès au contenu de ce livre.
2. Procédé selon la revendication 1, caractérisé en ce qu'il comprend, au titre des outils d'accès, une génération d'un index à partir des données sélectionnées dans un document, cet index signalant l'ensemble des pages dans lesquelles chacune des informations apparaît.
3. Procédé selon la revendication 2, caractérisé en ce que l'index est généré à partir d'informations surlignées.
4. Procédé selon l'une des revendications 1 ou 2, caractérisé en ce que l'index est généré également par des informations signalées par un traitement spécifique des caractères les composant.
5. Procédé selon l'une des revendications 2 à 4, caractérisé en ce que l'index comporte une zone pour saisir un terme à rechercher dans l'index.
6. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce qu'il comprend en outre une génération d'un résumé à partir d'informations sélectionnées au sein d'un document.

- 14 -

7. Procédé selon la revendication 6, caractérisé en ce que le résumé est généré à partir d'informations surlignées.

5

8. Procédé selon l'une des revendications 6 ou 7, caractérisé en ce que le résumé est généré à partir de notes prises dans d'autres documents d'un livre.

10 9. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce qu'on associe à un document au moins une note apposée sur ce document à la demande d'un utilisateur.

15 10. Procédé selon la revendication 9, caractérisé en ce que les notes se présentent sous la forme d'une feuille autocollante à apposer sur un document, dans laquelle l'utilisateur peut saisir une portion de texte.

20 11. Procédé selon l'une des revendications 2 à 5 et l'une des revendications 9 ou 10, caractérisé en ce que la recherche dans l'index peut porter aussi sur les notes apposées.

25 12. Procédé selon l'une des revendications 2 à 5 et l'une des revendications 9 ou 10, caractérisé en ce que la recherche dans l'index porte uniquement sur les notes apposées.

30 13. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce qu'il comprend en outre une génération d'un sommaire à partir de l'ensemble des documents consultés.

- 15 -

14. Procédé selon la revendication 13 , en ce que chaque document consulté est ajouté au sommaire de façon chronologique.
- 5 15. Procédé selon l'une des revendications 13 ou 14, caractérisé en ce qu'il comprend en outre une réorganisation du sommaire de façon hiérarchique.
- 10 16. Procédé selon la revendication 15, appliqué à des documents prélevés sur un réseau de communication tel qu'Internet, caractérisé en ce que la réorganisation du sommaire est réalisée suivant l'arborescence des adresses URL.
- 15 17. Procédé selon l'une des revendications 13 à 16, caractérisé en ce qu'il comprend en outre une possibilité pour un utilisateur de déplacer l'ensemble des liens du sommaire de façon à générer sa propre arborescence.
- 20 18. Procédé selon la revendication 17, caractérisé en ce qu'il comprend en outre un masquage d'une ou plusieurs branches de cette arborescence.
- 25 19. Procédé selon l'une quelconque des revendications 13 à 18, caractérisé en ce qu'il comprend en outre un signalement par le sommaire de documents modifiés sur des sites distants.
- 30 20. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce qu'il comprend en outre une génération d'un dossier contenant toutes les pages consultées.
- 35 21. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce qu'il comprend en outre

- 16 -

une attribution de durées de vie prédéterminées à des documents contenus dans un livre.

22. Système pour qualifier des données au sein de documents préalablement collectés, notamment en visitant des sites connectés à un réseau de communication, mettant en œuvre le procédé selon l'une quelconque des revendications précédentes, associé à des moyens pour collecter des documents, des moyens pour stocker ces documents collectés, des moyens pour traiter ces documents collectés et stockés et à des moyens pour visualiser et consulter le contenu de ces documents sous la forme de pages d'un livre, caractérisé en ce qu'il comprend des moyens pour sélectionner des données au sein des pages consultées.

23. Système selon la revendication 22, caractérisé en ce qu'il comprend en outre des moyens pour générer des outils d'accès au contenu des documents stockés au sein d'un livre, notamment un index et un résumé.

24. Système selon l'une des revendications 22 ou 23, caractérisé en ce qu'il comprend en outre des moyens pour générer un sommaire indiquant l'ensemble des documents consultés qui constituent un livre.

25. Application du procédé et du système de qualification de données selon l'une quelconque des revendications précédentes au rangement et à la consultation de documents collectés au cours d'une navigation sur un réseau de communication, notamment le réseau Internet.

1 / 5

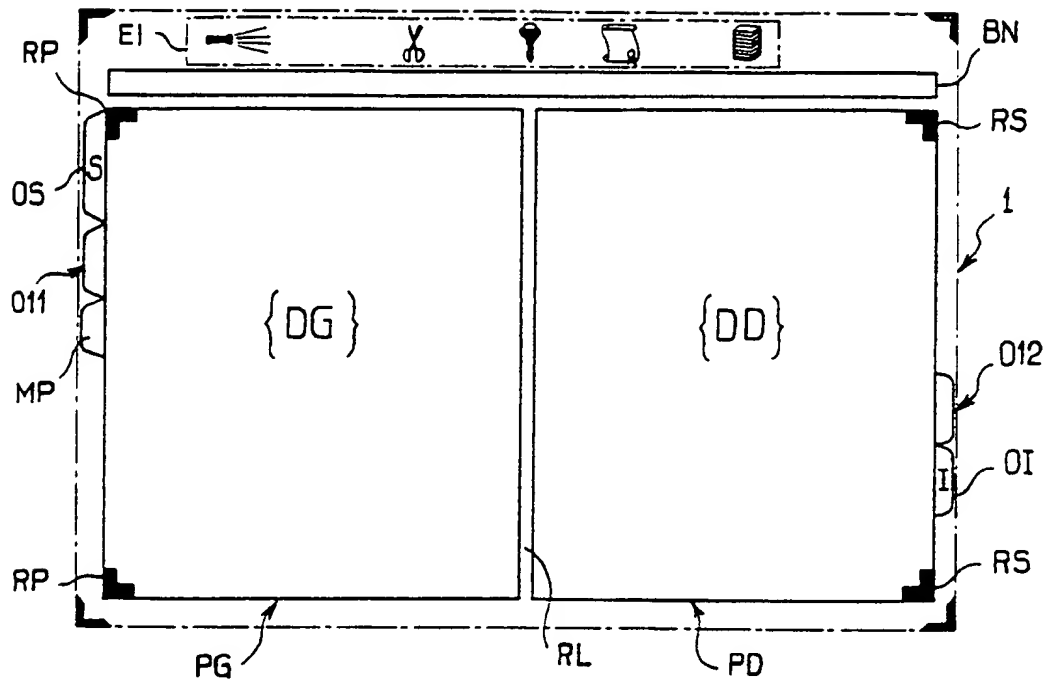


FIG. 1

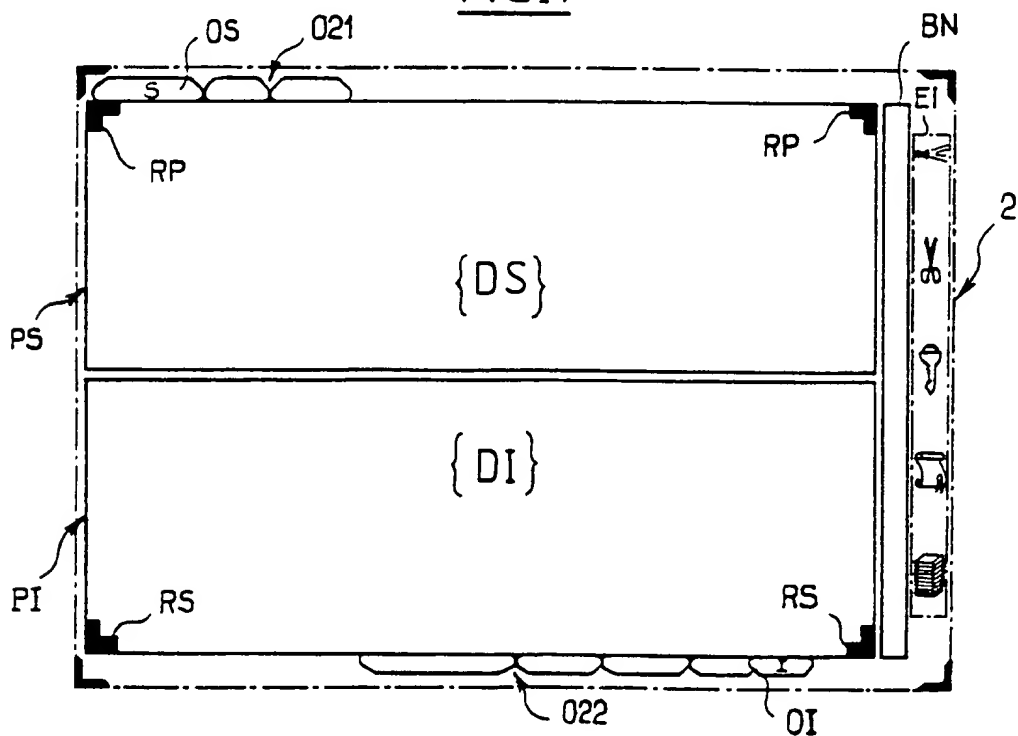


FIG. 2

2 / 5

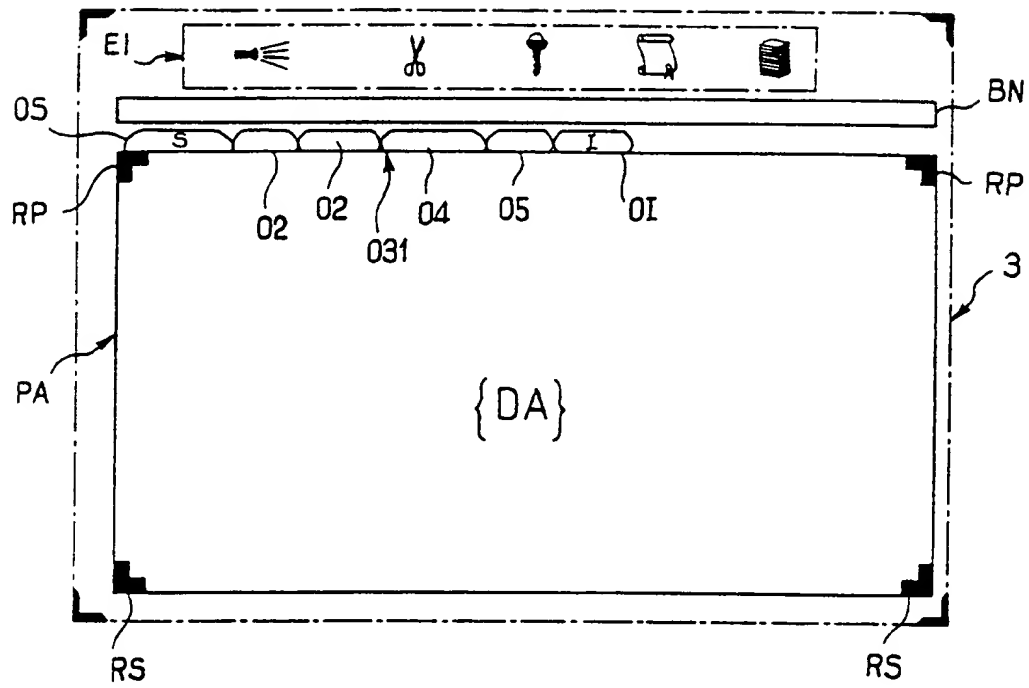


FIG. 3

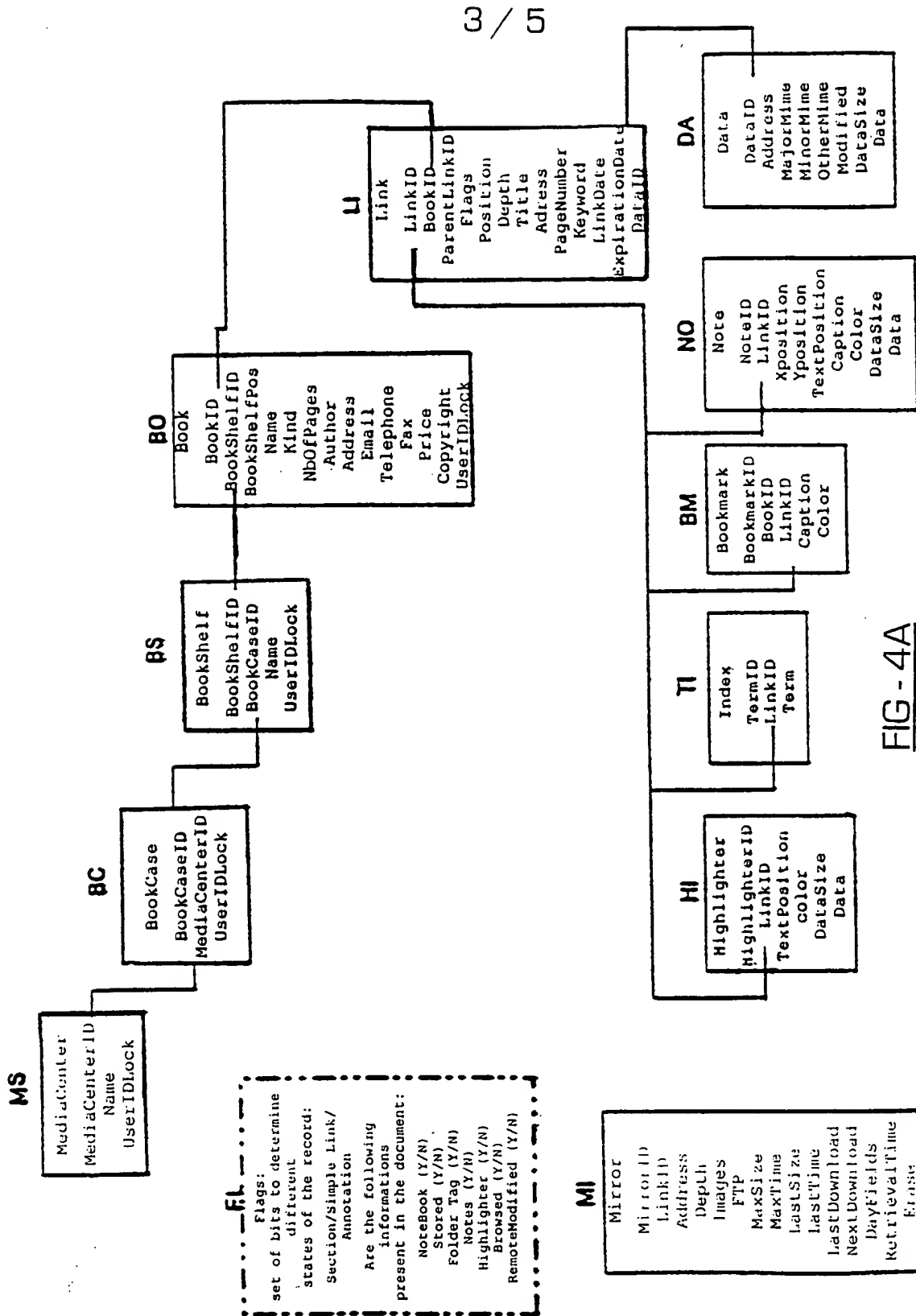


FIG - 4A

4 / 5

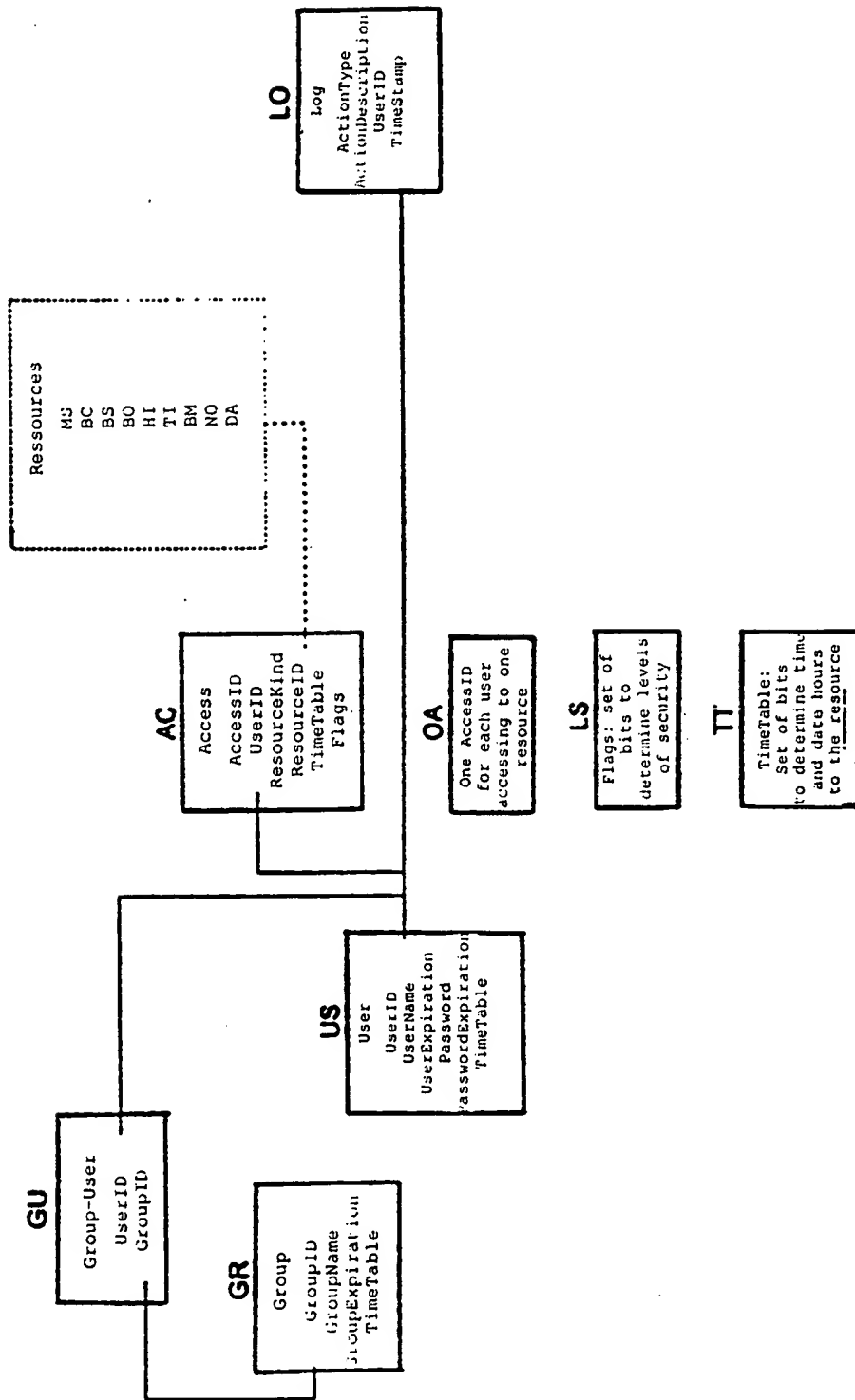


FIG - 4B

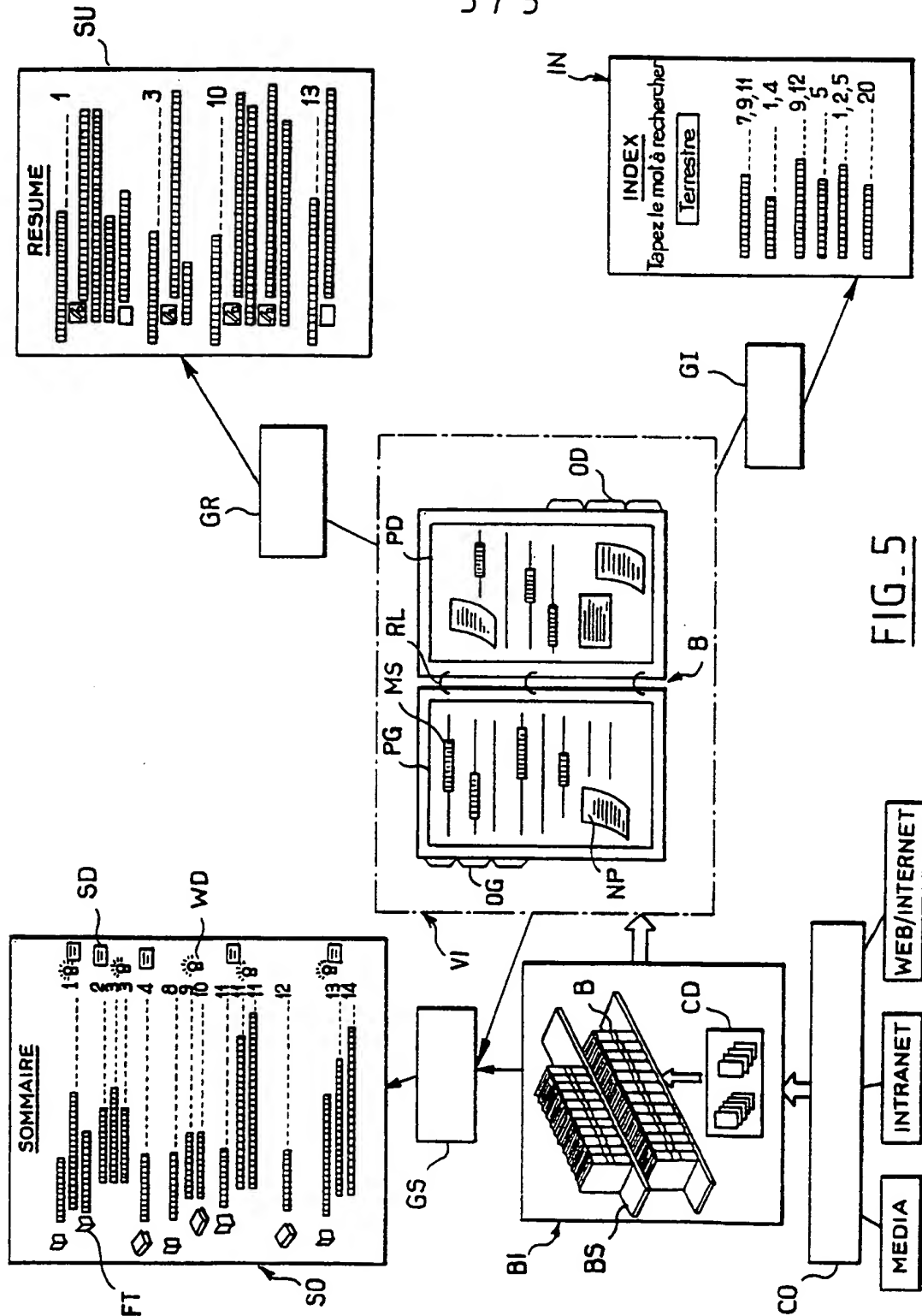


FIG. 5

INTERNATIONAL SEARCH REPORT

Interr. Application No

PCT/FR 97/00983

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 G06F3/033 G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 92 08199 A (GO CORP) 14 May 1992	1,2,5, 22,23
Y	see the whole document	25
A	<p>-----</p> <p>"BeOS Makes a Splash"</p> <p>PC MAGAZINE,</p> <p>vol. 16, no. 5, 4 March 1997,</p> <p>page 9 XP002054815</p> <p>see the whole document</p>	1,2,22
Y		25
A	<p>-----</p> <p>"SEARCHING FOR ANNOTATION IN DOCUMENTS ON A COMPUTER"</p> <p>RESEARCH DISCLOSURE,</p> <p>no. 318, 1 October 1990,</p> <p>page 816 XP000148656</p> <p>see the whole document</p> <p>-----</p>	1,9-12

-/--

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

6 February 1998

Date of mailing of the international search report

19/02/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Fournier, C

INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 97/00983

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>ROGERS E H: "NOTEMAKER: A HYPER-DESIGN NOTEBOOK TO SUPPORT INFORMAL COMMUNICATION IN THE ENGINEERING DESIGN PROCESS" ENGINEERED COMMUNICATION, ORLANDO, OCT. 30 - NOV. 1, 1991, vol. 1, 30 October 1991, INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, pages 141-144, XP000319134 see the whole document</p> <p>-----</p>	<p>1,2,9, 13,15, 16,22,25</p>

INTERNATIONAL SEARCH REPORT

information on patent family members

Intern. Application No

PCT/FR 97/00983

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9208199 A	14-05-92	AU 9031191 A	26-05-92
		CA 2093123 A	01-05-92
		EP 0555403 A	18-08-93
		JP 6502734 T	24-03-94

RAPPORT DE RECHERCHE INTERNATIONALE

Dem: Internationale No

PCT/FR 97/00983

A. CLASSEMENT DE L'OBJET DE LA DEMANDE
CIB 6 G06F3/033 G06F17/30

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)

CIB 6 G06F

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés)

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	WO 92 08199 A (GO CORP) 14 mai 1992	1,2,5, 22,23
Y	voir le document en entier	25
A	--- "BeOS Makes a Splash" PC MAGAZINE, vol. 16, no. 5, 4 mars 1997, page 9 XP002054815	1,2,22
Y	voir le document en entier	25
A	--- "SEARCHING FOR ANNOTATION IN DOCUMENTS ON A COMPUTER" RESEARCH DISCLOSURE, no. 318, 1 octobre 1990, page 816 XP000148656	1,9-12
	voir le document en entier	

	-/--	

☒ Voir la suite du cadre C pour la fin de la liste des documents

☒ Les documents de familles de brevets sont indiqués en annexe

* Catégories spéciales de documents cités:

"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent

"E" document antérieur, mais publié à la date de dépôt international ou après cette date

"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (elle qu'indiquée)

"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens

"P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

"S" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

6 février 1998

Date d'expédition du présent rapport de recherche internationale

19/02/1998

Nom et adresse postale de l'administration chargée de la recherche internationale

Office Européen des Brevets, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Fournier, C

RAPPORT DE RECHERCHE INTERNATIONALE

Dém. Internationale No
PCT/FR 97/00983

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités. avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	<p>ROGERS E H: "NOTEMAKER: A HYPER-DESIGN NOTEBOOK TO SUPPORT INFORMAL COMMUNICATION IN THE ENGINEERING DESIGN PROCESS" ENGINEERED COMMUNICATION, ORLANDO, OCT. 30 - NOV. 1, 1991, vol. 1, 30 octobre 1991, INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. pages 141-144, XP000319134 voir le document en entier -----</p>	<p>1,2,9, 13,15, 16,22,25</p>

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Dema internationale No

PCT/FR 97/00983

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 9208199 A	14-05-92	AU 9031191 A	26-05-92
		CA 2093123 A	01-05-92
		EP 0555403 A	18-08-93
		JP 6502734 T	24-03-94

PTO 03-3010

CY=WO DATE=19981210 KIND=A1
PN=98-55914

METHOD AND SYSTEM FOR DATA DESIGNATION IN PRE-COLLECTED DOCUMENTS
[PROCEDE ET SYSTEME DE QUALIFICATION DE DONNEES AU SEIN DE DOCUMENTS
PREALABLEMENT COLLECTES]

S. Dubois

UNITED STATES PATENT AND TRADEMARK OFFICE
Washington, D.C. MAY 2003

Translated by: FLS, Inc.

PUBLICATION COUNTRY	(10): France
DOCUMENT NUMBER	(11): WO 98/55914
DOCUMENT KIND	(12): International Patent Application
PUBLICATION DATE	(43): 19981210
PUBLICATION DATE	(45):
APPLICATION NUMBER	(21): PCT/FR97/00983
APPLICATION DATE	(22): 19970603
ADDITION TO	(61):
INTERNATIONAL CLASSIFICATION	(51): G06F 3/033
DOMESTIC CLASSIFICATION	(52):
PRIORITY COUNTRY	(33):
PRIORITY NUMBER	(31):
PRIORITY DATE	(32):
PRIORITY COUNTRY	(33):
PRIORITY NUMBER	(31):
PRIORITY DATE	(32):
PRIORITY DATE	(32):
DESIGNATED CONTRACTING STATES	(84): AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE
INVENTORS	(72): M. Sarars, F. Frachon, P. Sylvestre de Sacy Nougaret Durafort, T. Verrechia
APPLICANT	(71): Tereis Corporation
TITLE	(54): Procédé et système de qualification de données au sein de documents préalablement collectés
FOREIGN TITLE	(54A): Method and system for data designation in pre-collected documents

(54) Method and system for data designation in pre-collected documents

/0*

(57) ABSTRACT

The invention concerns a method for data designation in pre-collected documents, in particular from sites connected to a communication network, the said document being consulted in the form of pages of a book of which they comprise the content. Data or information can be selected within each page of a document during the retrieval process and contained in a book, in order to generate from this selected data access tools to the content of this book, especially an index, table of contents or summary. The method is useful especially for off-line retrieval of documents gathered on the web.

The present invention pertains to a method of designating data within pre-collected documents. It also envisages a system for its implementation.

/1

The development of information media and communication networks such as the Internet has led to a great increase of data and information that can be consulted and collected. The users of these new techniques are quickly drowned in volumes of data that one must rank, classify and designate, especially when one desires to make off-line consultations of these data. Presently, while experienced internet users can use fairly well personal techniques for classifying thousands of collected HTML pages, this is not true for new internet users who access a data universe that is unknown to them and who in any case do not have the skill and knack of the aforementioned internet users.

* Number in the margin indicates pagination in the foreign text.

The aim of the present invention is to propose a method of designating data that will allow simple and quick building of books from sets of documents collected from various information sources.

This is achieved with a method of designating within pre-collected documents, especially from sites connected to a communication network, these documents that can be consulted in the form of pages of a book of which they comprise the contents. One should note that there is already a software product called Lois (web navigator) that integrates a presentation of documents in the form of pages of a book. /2

According to the invention data can be selected within each page of a document during consultation and contained in a book, so as to generate from the said selected data some access tools to the contents of this book.

One thus obtains with the method of designation according to the invention a means of organization and management of data that goes well beyond a simple compilation of consulted pages since it becomes possible to process all the data contained in the documents comprising a book in a global form thanks to some access tools configured for each book.

For example, as access tools the method according to the invention can include generation of an index from data presented in a document, this index indicating all of the pages in which the information appears. This index can be generated in particular from

highlighted information, but also from data indicated by specific treatment of the characters comprising it. It can include an area for capturing a term to be searched in the index.

An access tool to the contents of the documents consists in a summary generated from information selected within a document. This summary can also be supplied from notes taken in other documents of a book.

The designation method according to the invention can also advantageously provide for combining with a document one or several notes appended to this document at the request of a user. These notes are present for example in the form of a self-stick sheet to be attached to a document, in which the user can capture a text /3 portion. These notes can have the property of being transparent, thus allowing one to visualize the contents of the note without having to cover up the document. The search in the index can also pertain to these attached notes.

In a preferred version of the method according to the invention a table of contents is furnished from the set of consulted documents. Each consulted document is thus added to the table of contents chronologically, a new organization of the table of contents in hierarchical manner thus being envisaged. In particular, when one is concerned with collected documents on a communication network such as the Internet the re-organization of the table of contents can be done based on the tree diagram of the URL addresses.

The method according to the invention can additionally advantageously include the possibility for the user to move the set of links of the table of contents so as to generate his own tree diagram, as well as a masking of one or several branches of this tree diagram.

A description by the table of contents of modified documents to remote sites, of locally protected documents or of designated documents (annotations, highlighted passages in the document), can also be provided. In practice the method according to the invention includes moreover the generation of a report containing all the consulted pages.

According to another feature of the invention a system is proposed for designating data within pre-selected documents, especially by visiting sites connected to a communication network, implementing the method according to the invention, combined with some means for collecting documents, means for processing these collected documents and stored and with some means for visualizing and consulting the content of these documents in the form of pages of a book, characterized in that it includes means for selecting data within the consulted pages. This system also includes means for generating access tools to the contents of the documents stored in a book, especially an index, a table of contents and a summary. /4

A personal blank page can be inserted in the current position in order to gather text or graphic notes or any other element. One can also provide the possibility of using PAO (publication) software for this personal page.

Other particular features and advantages of the invention will be apparent from the following description. In the attached drawings that are given as non-limiting examples:

- Figure 1 shows a first example of visualization of documents in "book" format by the method according to the invention;
- Figure 2 shows a second example of visualization of documents in "notebook" format by the method according to the invention;
- Figure 3 shows a third example of visualization of documents in "block notes" format by the method according to the invention;
- Figure 4A shows the structure of a data base used with the method according to the invention;
- Figure 5 is a block diagram view of the essential functions achieved by the method according to the invention.

We shall now describe some examples of document visualization obtained with the method of designation according to the invention, with reference to Figs. 1 to 3.

Screen **1** includes two left and right side pages PG, PD respectively that are joined to a binding RL, and a set of icons arranged for example on the upper part of the screen. The book B1 thus visualized includes a first set **011** of tabs arranged on its left side and a second set **012** of tabs arranged on its right side. It is obvious that as the pages are turned the respective numbers of tabs located instantaneously on the left and right respectively are variable. The first set **011** of tabs includes a tab OS "table of contents" located preferably at the top of this first set of tabs and a page marker MP, while the second set **012** of tabs includes a tab **01** "Index" located preferably at the bottom of this second set of tabs. The pages of the left and right PG, PD receive respectively active documents DG, DD. The screen **1** includes also a navigation strip BN located for example on its upper part and a set of command icons E1. Shortened zones RP, RS are also provided for reaching the next and preceding pages respectively.

In the second example of visualization shown by Fig. 2 the screen **2** includes a first upper page PS and a second lower page P1 that each contain some active documents DS, DI in course of visualization. The notebook B2 thus visualized includes a first set of tabs **021** located on the upper edge of the upper part PS and a second set of tabs **022** located on the lower edge of the lower part P1. The set of icons EI is for example arranged on the right side of

screen 2. The first set of tabs **021** includes for example a "table of contents" tab OS, while the second set of tabs **022** includes for example a tab "index" OI.

/6

One can also use a mode of visualization of the block notes kind with reference to Fig. 3, in which a screen 3 includes a set of tabs **031** arranged for example on the upper edge of a single page PA that contains an active document DA in the course of visualization. The set of tabs **031** includes for example a tab "table of contents" OS, some interleaved tabs **02-05** and an index tab **0I**.

A database DB used with the method according to the invention includes, with reference to Figs. 4A and 4B, a documentation center MS, a bookcase BC, some shelves BS, some books BO, some indexed terms TI, some links LI, some data DA, some bookmarks BM, some notes NO and some highlighters HI. It also includes some flags FL, an access card catalog AC, some flags LS for determining security levels, a timetable TT, an event log LO, some users US, some user groups GR, some relationships between groups and users GU and a unique access identifier OA.

The media or documentation center MS, called also later the MediaCenter suite, is the logical paradigm of a documentation center. It allows one to identify the large information sources. As a non-limiting example, within a company, it is possible to install several data base servers that can be consulted indiscriminately. The

MediaCenter gathers together the set of bookcases called later BookCase.

The BookCase is a storage space for a large number of books grouped logically by a user. It is the furniture paradigm that a bookcase represents. It is associated with a unique MediaCenter.

The BookShelf is the paradigm of a bookshelf. It allows one 17 to gather thematically a set of books. There is no logical limit to the number of books. Only the representation on the screen of a shelf in the bookcase can limit their number.

The books are arranged in shelves and can be moved in any of the aforementioned hierarchical elements.

The MediaCenters can be some sharable elements on the Web. A user navigating on the Internet can decide to consult a library of information. He can thus go from one MediaCenter into another MediaCenter.

The set on indexed words can be regrouped in the form of a general book of indexes in order to allow a search on the sets of books.

We shall now describe some examples of management of the rights of access to information as part of the method according to the invention. We shall define a set of services that allow the administration of access rights to information.

Each document has a list of readers and groups of readers having specific rights to this document. The sources have rights attributed to users. Some rights can be associated with:

- Documentation centers (MediaCenter);
- Bookcases (BookCase);
- Groups of books or shelves (Bookshelf);
- Books,
- Sections or chapters;
- Documents (or links) of a book.

The following rights can be applied:

- Reading;
- Writing; /8
- Adding (one book for groups of books, one document for books);
- Appropriation (one group of books, one book, a document),
- Backup;
- Time ranges (Timetable).

A user authorized to prohibit a group of books cannot however prohibit the documents of a book, unless he is explicitly given this right.

The ownership of an element does not imply the set of rights available but the possibility of modifying any right whatsoever of this element. An administrator has a right of appropriation to all the elements of the library. The idea of ownership allows one to identify any user who is appropriate for a document.

The right of backup does not imply the right of reading: a specific operator can hold this right without necessarily being authorized to consult the data.

The clocks allow one to limit access to information, for example by analogy with the open hours of a library.

Only the operators created by the administrator have the opportunity to create readers and groups of readers. One can define for operators the following rights:

- Adding users;
- Restricting users;
- Restricting readers that they have created;
- Restricting groups of readers that they have created;
- Restricting readers;
- Restricting groups of readers.

There is an administrative group whose members have the same rights as the administrator. The administrator can never be restricted.

19

One can also attribute a lifetime to the data that is propagated through the field ExpirationDate of the database. The insertion of this lifetime allows one to solve problems of confidentiality, storage capacity, and preservation of pertinent data. It reinforces also the paradigm of the book since the readers are accustomed to read dailies, weeklies, and monthlies.

The users and the user groups also have a lifetime that allow one to restrict for a certain time the rights of some users. In addition, it is possible to set an expiration date of the passwords of the users to force them to renew their passwords and thus increase security.

We shall now describe the essential functions obtained by the method according to the invention, with reference to Fig. 5. Information is collected (CO) on the Web/Internet, on an Intranet network or on some information media such as CD-ROM. This information or documents are then classified (DC) in order to be stored in books B arranged on shelves BS, within a library BI. When a user wishes to consult a document, the book B containing this document is retrieved from the library and is opened in order to allow a visualization VI of the pages of this book. Thus, one can visualize a left page PG and a right page PD that can turn about a binding RL, these pages being possibly with tabs OG, OD. On the pages of the documents the user can at his liking highlight (MS) some words or excerpts of the document, or attach some notes NP.

/10

During the building of the book the method according to the invention includes the following functions:

- Generation GS of a table of contents SO;
- Generation GR of a summary SU;
- Generation GI of an index IN.

The highlighter has the function of allowing the highlighting of information of information by the user. Thus, when consulting a document again he can see more easily what he is interested in. This function lets one also share key information with other users, especially in an Intranet architecture. The highlighted words are also used for the creation of an index.

The summary is a document generated from highlighted elements and notes made in other documents of the book.

The table of contents is generated from the set of consulted documents. Each consulted document is added to the table of contents chronologically. A user can have access to the table of contents at any time of the consultation. He can request reorganization of the table of contents hierarchically, for example according to the tree diagram of the URL addresses in the case of documents obtained on the Internet. The set of links of the table of contents can be moved by the user so as to generate its own tree diagram. Any branch of this tree diagram can be masked. The table of contents is capable of indicating to the user the modified documents on remote sites. A

special report called "unclassified pages" contains all the consulted pages.

Each remotely loaded document and belonging to the table of contents is analyzed in order to extract some key words from it. These key words let one generate an index. They are emphasized in a document by some character functions (bold, italics, ...) or by /11 using the highlighting function.

The index indicates the set of pages where each of the key words appears, even if it was emphasized in a single document.

The index makes available a capture zone that lets one move instantaneously to the search phrase in the index. This capture zone can also allow a more complete search in the entire book. In this case a new document containing the result of the search is added to the book.

A user can associate with a document one or several notes that are positioned to the liking of the user on the document. During a search in the index, it is possible to specify that the search is being done also in the notes, or only in the notes. The notes are presented in the form of a self-attaching sheet in which the user can capture a text portion.

One should note that the designation method according to the invention can be combined with any publishing and printing method to physically produce true books, notebooks or notepads or any other real medium. In addition the designation method according to the

invention can consider any application available in the multimedia field.

Of course the invention is not limited to the examples that have just been described and many variations can be made to these examples without departing from the scope of the invention. Thus, other tools that those just described can be employed for the designation of data. One should note that the data and documents in question by the present invention can be of any origin whatsoever and might have been collected on open communication networks such as the Internet or /12 private one (Intranet), or even on some information media of any kind (CD-ROM, DVD-ROM, and so forth).

CLAIMS

/13

1. Method for designating data within pre-collected documents, especially from sites connected to a communication network, these documents capable of being consulted in the form of pages of a book of which they comprise the contents, characterized in that the data can be selected within each page of a document in the course of consultation and contained in a book, so as to generate from the said selected data access tools to the contents of this book.

2. Method according to Claim 1 characterized in that it includes as access tools the generation of an index from data selected in a document, this index indicating the set of pages in which each of the information appears.

3. Method according to Claim 2 characterized in that the index is generated from highlighted information.

4. Method according to Claims 1 or 2 characterized in that the index is generated also by some information indicated by specific treatment of the characters that make it up.

5. Method according to one of the Claims 2 to 4 characterized in that the index includes a zone for capturing a phrase to search in the index.

6. Method according to any of the preceding claims characterized in that it includes in addition the generation of a summary from information selected in a document.

7. Method according to Claim 6 characterized in that the /14 summary is generated from highlighted information.

8. Method according to one of the Claims 6 or 7 characterized in that the summary is generated from notes taken in other documents of a book.

9. Method according to any of the preceding claims characterized in that one associated with a document at least one note attached to this document at the request of a user.

10. Method according to Claim 9 characterized in that the notes are present in the form of a self-attaching sheet to be posted on a document in which the user can capture a text portion.

11. Method according to one of the Claims 2 to 5 and one of the claims 9 or 10 characterized in that the search in the index can pertain also to the posted notes.

12. Method according to one of the Claims 2 to 5 and one of the claims 9 or 10 characterized in that the search in the index pertains only to the posted notes.

13. Method according any of the preceding claims characterized in that it includes in addition the generation of a table of contents from the set of consulted documents:

14. Method according to Claim 13 characterized in that each /15 consulted document is added to the table of contents chronologically.

15. Method according to one of the Claims 13 or 14 characterized in that it includes in addition a reorganization of the table of contents hierarchically.

16. Method according to Claim 15 applied to documents obtained in a communication network such as the Internet, characterized in that the reorganization of the table of contents is done according to the tree diagram of the URL addresses.

17. Method according to one of the Claims 13 to 16 characterized in that it includes in addition the possibility for a user to move the set of links from the table of contents so as to generate his own tree diagram.

18. Method according to Claim 17 characterized in that it includes in addition a masking of one or several branches of this tree diagram.

19. Method according any of the Claims 13 to 18 characterized in that it includes in addition reference by the table of contents of modified documents to some remote sites.

20. Method according to any of the preceding claims characterized in that it includes in addition generation of a report containing all the consulted pages.

21. Method according to any of the preceding claims characterized in that it includes in addition an attribution of predetermined lifetimes to some documents contained in a book. /16

22. System for designating data within pre-collected documents, especially by visiting sites connected to a communication network, using the method according to the invention based on any of the preceding claims, combined with some means for collecting documents, means for storing these collected documents, some means for processing these collected and stored documents and some means for visualizing and consulting the contents of these documents in the form of pages of a book, characterized in that it includes some means for selecting data within the consulted pages.

23. System according to Claim 22 characterized in that it includes in addition some means for generating access tools to the

contents of the documents stored in a book, especially an index and a summary.

24. System according to one of the Claims 22 or 23 characterized in that it includes in addition some means for generating a table of contents indicating the set of consulted documents that comprise a book.

25. Application of the method and system of data designation according to any of the preceding claims for orderly arrangement and consultation of documents collected in the course of navigation on a communication network, especially the Internet network.

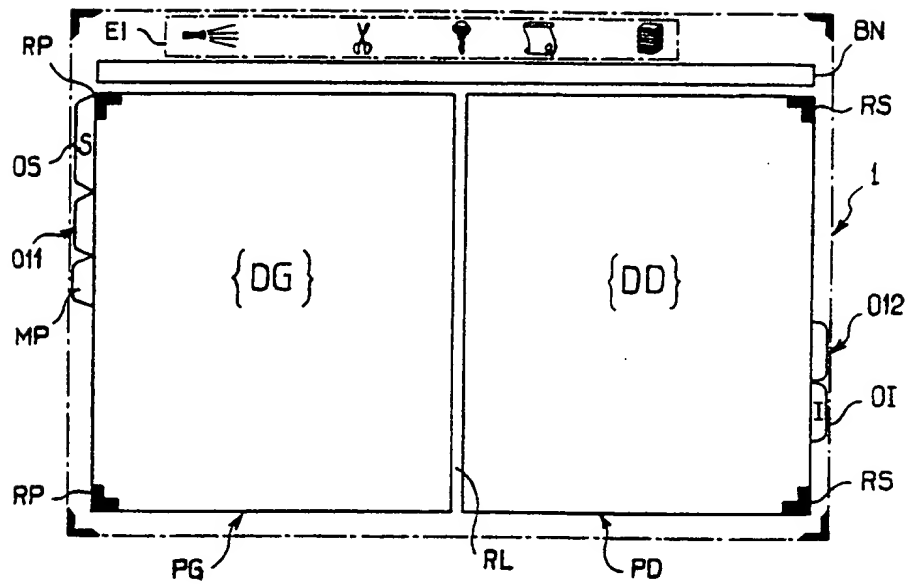


FIG. 1

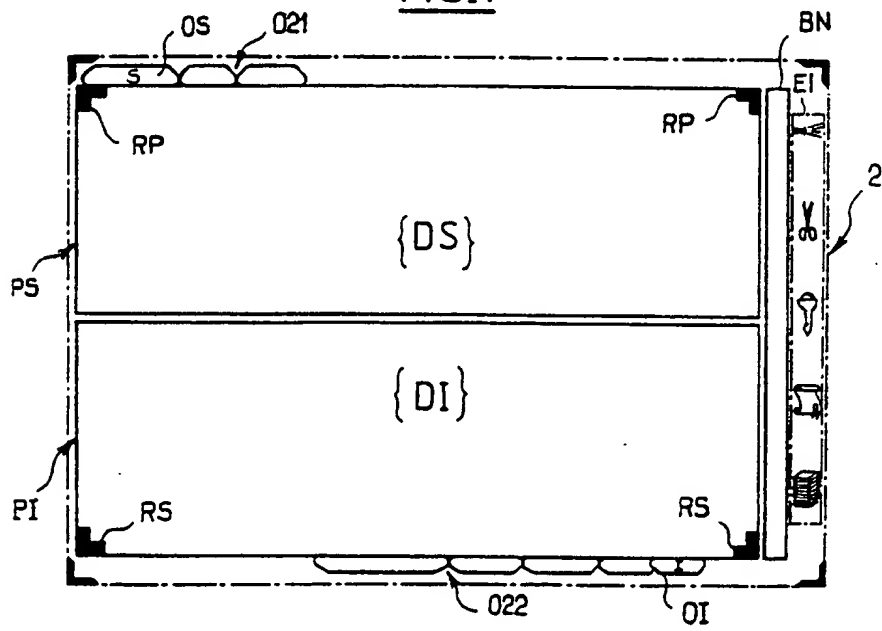


FIG. 2

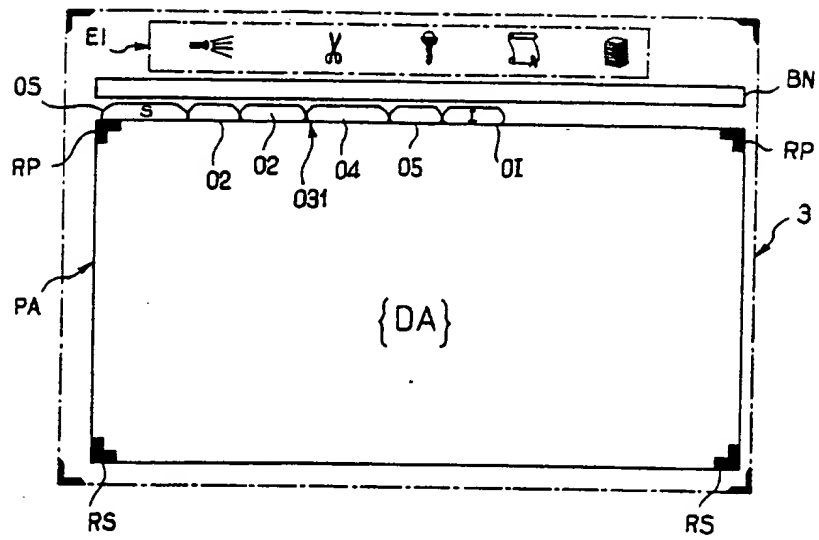


FIG. 3

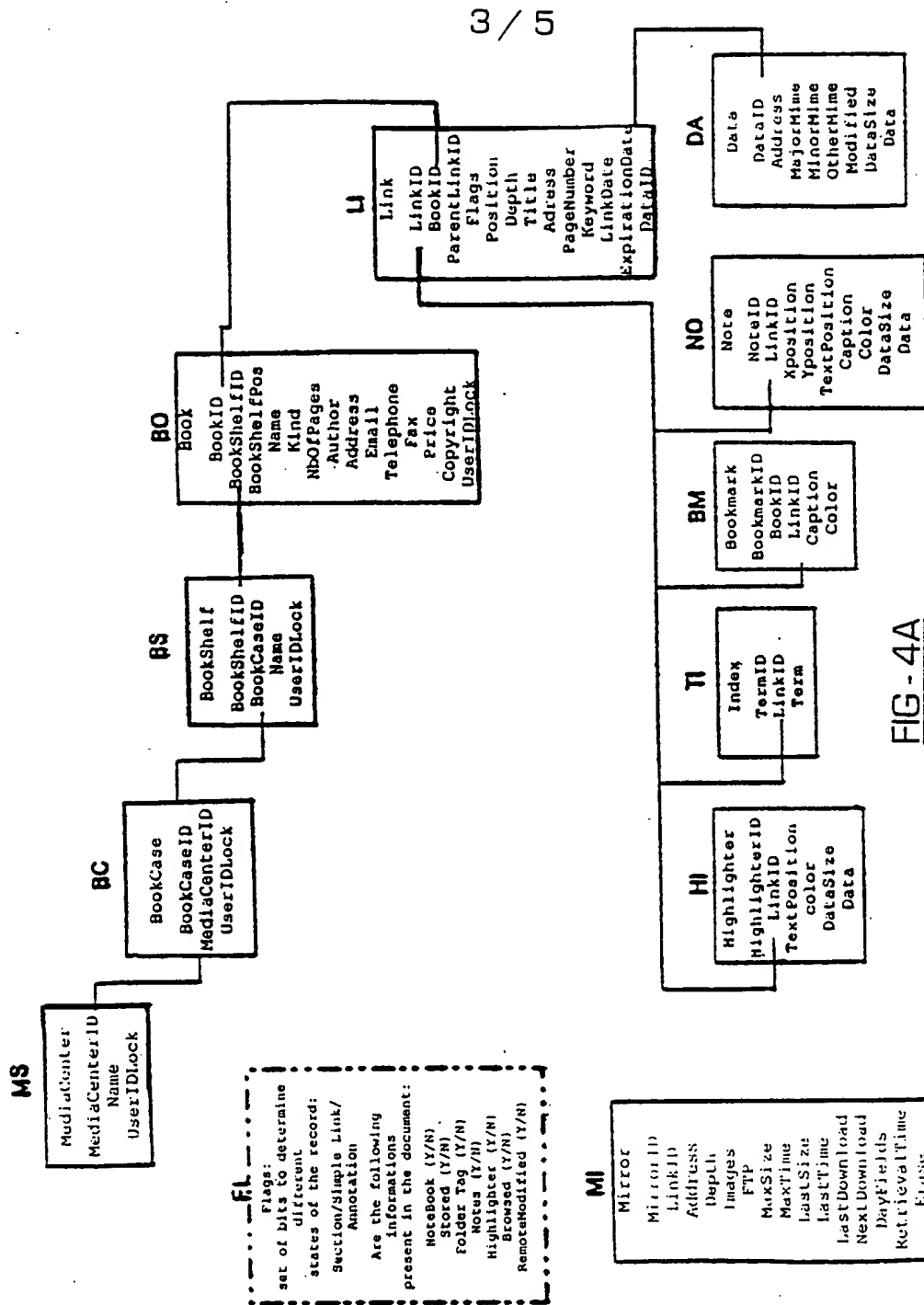


FIG - 4A

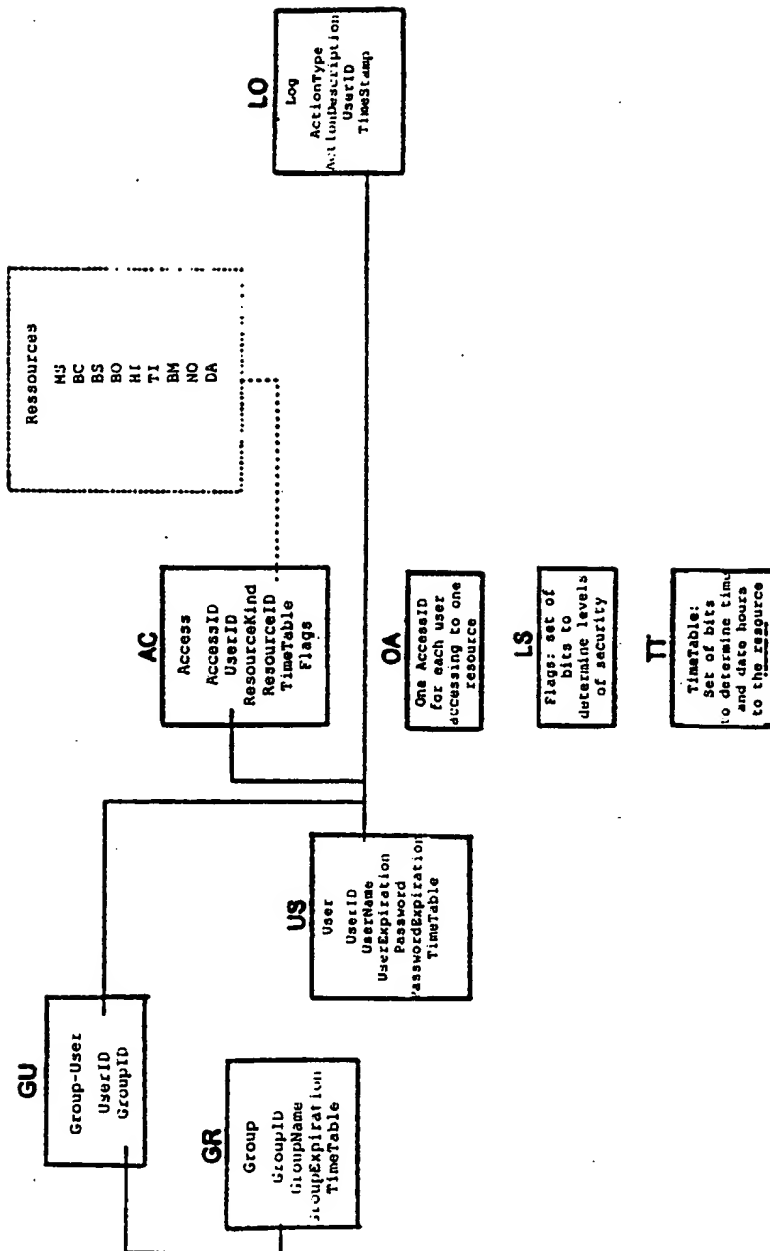


FIG - 4B

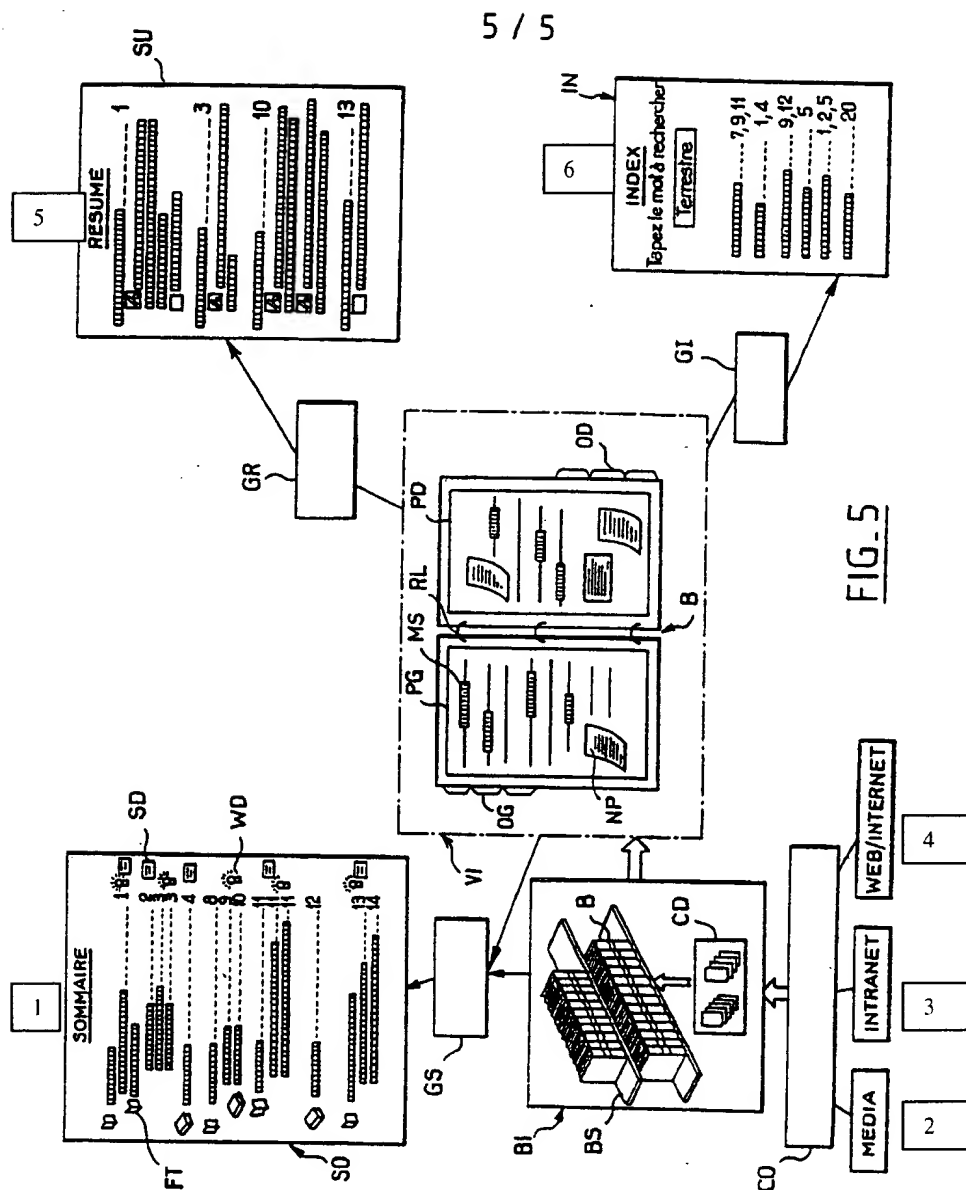


Figure 5. Key: 1-table of contents; 2-media; 3-intranet; 4-web/internet; 5-summary; 6-index, type in the word to be searched, terrestrial.